

Research Article

A Multifactor Extension of Linear Discriminant Analysis for Face Recognition under Varying Pose and Illumination

Sung Won Park and Marios Savvides

Electrical and Computer Engineering Department, Carnegie Mellon University, 5000 Forbes Avenue Pittsburgh, PA 15213, USA

Correspondence should be addressed to Sung Won Park, sungwonp@cmu.edu

Received 11 December 2009; Revised 27 April 2010; Accepted 20 May 2010

Academic Editor: Robert W. Ives

Copyright © 2010 S. W. Park and M. Savvides. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Linear Discriminant Analysis (LDA) and Multilinear Principal Component Analysis (MPCA) are leading subspace methods for achieving dimension reduction based on supervised learning. Both LDA and MPCA use class labels of data samples to calculate subspaces onto which these samples are projected. Furthermore, both methods have been successfully applied to face recognition. Although LDA and MPCA share common goals and methodologies, in previous research they have been applied separately and independently. In this paper, we propose an extension of LDA to multiple factor frameworks. Our proposed method, Multifactor Discriminant Analysis, aims to obtain multilinear projections that maximize the between-class scatter while minimizing the withinclass scatter, which is the same core fundamental objective of LDA. Moreover, Multifactor Discriminant Analysis (MDA), like MPCA, uses multifactor analysis and calculates subject parameters that represent the characteristics of subjects and are invariant to other changes, such as viewpoints or lighting conditions. In this way, our proposed MDA combines the best virtues of both LDA and MPCA for face recognition.

1. Introduction

Face recognition has significant applications for defense and national security. However, today, face recognition remains challenging because of large variations in facial image appearance due to multiple factors including facial feature variations among different subjects, viewpoints, lighting conditions, and facial expressions. Thus, there is great demand to develop robust face recognition methods that can recognize a subject's identity from a face image in the presence of such variations. Dimensionality reduction techniques are common approaches applied to face recognition not only to increase efficiency of matching and compact representation, but, more importantly, to highlight the important characteristics of each face image that provide discrimination. In particular, dimension reduction methods based on supervised learning have been proposed and commonly used in the following manner. Given a set of face images with class labels, dimension reduction methods based on supervised learning make full use of class labels of these images to learn each subject's identity. Then, a generalization of this dimension reduction is achieved for unlabeled test

images, also called out-of-sample images. Finally, these test images are classified with respect to different subjects, and the classification accuracy is computed to evaluate the effectiveness of the discrimination.

Multilinear Principal Component Analysis (MPCA) [1, 2] and Linear Discriminant Analysis (LDA) [3, 4] are two of the most widely used dimension reduction methods for face recognition. Unlike traditional PCA, both MPCA and LDA are based on supervised learning that makes use of given class labels. Furthermore, both MPCA and LDA are subspace projection methods that calculate low-dimensional projections of data samples onto these trained subspaces. Although LDA and MPCA have different ways of calculating these subspaces, they have a common objective function which utilizes a subject's individual facial appearance variations.

MPCA is a multilinear extension of Principal Component Analysis (PCA) [5] that analyzes the interaction between multiple factors utilizing a tensor framework. The basic methodology of PCA is to calculate projections of data samples onto the linear subspace spanned by the principal directions with the largest variance. In other words, PCA finds the projections that best represent the data. While PCA

calculates one type of low-dimensional projection vector for each face image, MPCA can obtain multiple types of low-dimensional projection vectors; each vector parameterizes a different factor of variations such as a subject's identity, viewpoint, and lighting feature spaces. MPCA establishes multiple dimensions based on multiple factors and then computes multiple linear subspaces representing multiple varying factors.

In this paper, we separately address the advantages and disadvantages of multifactor analysis and discriminant analysis and propose Multifactor Discriminant Analysis (MDA) by synthesizing both methods. MDA can be thought of as an extension of LDA to multiple factor frameworks providing both multifactor analysis and discriminant analysis. LDA and MPCA have different advantages and disadvantages, which result from the fact that each method assumes different characteristics for data distributions. LDA can analyze clusters distributed in a global data space based on the assumption that the samples of each class approximately create a Gaussian distribution. On the other hand, MPCA can analyze the locally repeated distributions which are caused by varying one factor under fixed other factors. Based on synthesizing both LDA and MPCA, our proposed MDA can capture both global and local distributions caused by a group of subjects.

Similar to our MDA, the Multilinear Discriminant Analysis proposed in [6] applies both tensor frameworks and LDA to face recognition. Our method aims to analyze multiple factors such as subjects' identities and lighting conditions in a set of vectored images. On the other hand, [6] is designed to analyze multidimensional images with a single factor, that is, subjects' identities. In [6], each face image constructs an n -mode tensor, and the low-dimensional representation of this original tensor is calculated as another n -mode tensor with a smaller size. For example, if we simply use 2-mode tensors, that is, matrices, representing 2D images, the method proposed in [6] reduces each dimension of the rows and columns by capturing the repeated tendencies in rows and the repeated tendencies in columns. On the other hand, our proposed MDA analyzes the repeated tendencies caused by varying each factor in a subspace obtained by LDA. The goal of MDA is to reduce the impacts of environmental conditions, such as viewpoint and lighting, from the low-dimensional representations obtained by LDA. While [6] obtains a single tensor with a smaller size for each image tensor, our proposed MDA obtains multiple low-dimensional vectors, for each image vector, which decompose and parameterize the impacts of multiple factors. Thus, for each image, while the low-dimensional representation obtained by [6] is still influenced by variance in environmental factors, multiple parameters obtained by our MDA are expected to be independent from each other. The extension of [6] to multiple factor frameworks cannot be simply drawn because this method is formulated only using a single factor, that is to say, subjects' identities. On the other hand, our proposed MDA decomposes the low-dimensional representations obtained by LDA into multiple types of factor-specific parameters such as subject parameters.

The remainder of this paper is organized as follows. Section 2 reviews subspace methods from which the proposed method is derived. Section 3 first addresses the advantages and disadvantages of multifactor analysis and discriminant analysis individually, and then Section 4 proposes MDA with the combined virtues of both methods. Experimental results for face recognition in Section 5 show that the proposed MDA outperforms major dimension reduction methods on the CMU PIE database and the Extended Yale B database. Section 6 summarizes the results and conclusions of our proposed method.

2. Review of Subspace Projection Methods

In this section, we review MPCA and LDA, two methods on which our proposed Multifactor Discriminant Analysis is based.

2.1. Multilinear PCA. Multilinear Principal Component Analysis (MPCA) [1, 2] is a multilinear extension of PCA. MPCA computes a linear subspace representing the variance of data due to the variation of each factor as well as the linear subspace of the image space itself. In this paper, we consider three factors: different subjects, viewpoints (i.e., pose types), and lighting conditions (i.e., illumination). While PCA is based on Singular Value Decomposition (SVD) [7], MPCA is based on High-Order Singular Value Decomposition (HOSVD) [8], which is a multidimensional extension of SVD.

Let \mathbf{X} be the $m_p \times n$ data matrix whose columns are vectored training images $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n$ with n_p pixels. We assume that these data samples are centered at zero. By SVD, the matrix \mathbf{X} can be decomposed into three matrices \mathbf{U} , \mathbf{S} , and \mathbf{V} :

$$\mathbf{X} = \mathbf{U}\mathbf{S}\mathbf{V}^T. \quad (1)$$

If we keep only the $m < n$ column vectors of \mathbf{U} and \mathbf{V} corresponding to the m largest singular values and discard the rests of the matrices, the sizes of the matrices in (1) are as follows: $\mathbf{U} \in \mathbb{R}^{n_p \times m}$, $\mathbf{S} \in \mathbb{R}^{m \times m}$, and $\mathbf{V} \in \mathbb{R}^{n \times m}$. For a sample \mathbf{x} , PCA obtains an m -dimensional representation:

$$\mathbf{y}_{\text{PCA}} = \mathbf{U}^T \mathbf{x}. \quad (2)$$

Note that these low-dimensional projections preserve the dot products of training images. We define the matrix $\mathbf{Y}_{\text{PCA}} \in \mathbb{R}^{m \times n}$ consisting of these projections obtained by PCA:

$$\mathbf{Y}_{\text{PCA}} = \mathbf{U}^T \mathbf{X} = \mathbf{S}\mathbf{V}^T. \quad (3)$$

Then, we can see that the Gram matrices of \mathbf{X} and \mathbf{Y}_{PCA} are identical since

$$\mathbf{G} = \mathbf{X}^T \mathbf{X} = \mathbf{Y}_{\text{PCA}}^T \mathbf{Y}_{\text{PCA}} = \mathbf{V}\mathbf{S}^2\mathbf{V}^T. \quad (4)$$

Since a Gram matrix is a matrix of all possible dot products, a set of \mathbf{y}_{PCA} also preserves the dot products of original training images.

While PCA parameterizes a sample \mathbf{x} with one low-dimensional vector \mathbf{y} , MPCA [1] parameterizes the sample using multiple vectors associated with multiple factors of a data set. In this paper, we consider three factors of face images: n_s identities (or subjects), n_v poses, and n_l lighting conditions. $\mathbf{x}_{i,p,l}$ denotes a vectored training image of the i th subject in the p th pose and the l th lighting condition. These training images are sorted in a specific order so as to construct a data matrix $\mathbf{X} \in \mathbb{R}^{m \times n_s n_v n_l}$:

$$\mathbf{X} = [\mathbf{x}_{1,1,1}, \mathbf{x}_{2,1,1}, \dots, \mathbf{x}_{n_s,1,1}, \mathbf{x}_{1,2,1}, \dots, \mathbf{x}_{n_s,n_v,n_l}]. \quad (5)$$

Using MPCA, an arbitrary image \mathbf{x} and a data matrix \mathbf{X} are represented as

$$\mathbf{x} = \mathbf{U}\mathbf{Z}(\mathbf{v}^{\text{subj}} \otimes \mathbf{v}^{\text{view}} \otimes \mathbf{v}^{\text{light}}), \quad (6)$$

$$\mathbf{X} = \mathbf{U}\mathbf{Z}(\mathbf{V}^{\text{subj}} \otimes \mathbf{V}^{\text{view}} \otimes \mathbf{V}^{\text{light}})^T, \quad (7)$$

respectively, where \otimes denotes the Kronecker product and \mathbf{U} is identical to the matrix \mathbf{U} in (1). A matrix \mathbf{Z} results from the pixel-mode flattening of a core tensor [1]. In (6), we can see that MPCA parameterizes a single image \mathbf{x} using three parameters: subject parameter $\mathbf{v}^{\text{subj}} \in \mathbb{R}^{n'_s}$, viewpoint parameter $\mathbf{v}^{\text{view}} \in \mathbb{R}^{n'_v}$, and lighting parameter $\mathbf{v}^{\text{light}} \in \mathbb{R}^{n'_l}$, where $n'_s \leq n_s$, $n'_v \leq n_v$, and $n'_l \leq n_l$. Similarly, \mathbf{X} in (7) is represented by three orthogonal matrices $\mathbf{V}^{\text{subj}} \in \mathbb{R}^{n_s \times n'_s}$, $\mathbf{V}^{\text{view}} \in \mathbb{R}^{n_v \times n'_v}$, and $\mathbf{V}^{\text{light}} \in \mathbb{R}^{n_l \times n'_l}$. The columns of each matrix span the linear subspace of the data space formed by varying each factor. Therefore, \mathbf{V}^{subj} , \mathbf{V}^{view} , and $\mathbf{V}^{\text{light}}$ consist of eigenvectors corresponding to the largest eigenvalues of three Gram-like matrices \mathbf{G}^{subj} , \mathbf{G}^{view} , and $\mathbf{G}^{\text{light}}$ respectively, where the (r, c) entry of these matrices is calculated as

$$\begin{aligned} \mathbf{G}_{rc}^{\text{subj}} &= \frac{1}{n_v n_l} \sum_{p=1}^{n_v} \sum_{l=1}^{n_l} \mathbf{x}_{r,p,l}^T \mathbf{x}_{c,p,l}, \\ \mathbf{G}_{rc}^{\text{view}} &= \frac{1}{n_s n_l} \sum_{i=1}^{n_s} \sum_{l=1}^{n_l} \mathbf{x}_{i,r,l}^T \mathbf{x}_{i,c,l}, \\ \mathbf{G}_{rc}^{\text{light}} &= \frac{1}{n_s n_v} \sum_{i=1}^{n_s} \sum_{p=1}^{n_v} \mathbf{x}_{i,p,r}^T \mathbf{x}_{i,p,c}. \end{aligned} \quad (8)$$

These three Gram-like matrices \mathbf{G}^{subj} , \mathbf{G}^{view} , $\mathbf{G}^{\text{light}}$, represent similarities between different subjects, different poses, and different lighting conditions, respectively. For example, \mathbf{G}^{subj} can be thought of as the average similarity, measured by the dot product, between the r th subject's face images and the c th subject's face images under varying viewpoints and lighting conditions.

Three orthogonal matrices \mathbf{V}^{subj} , \mathbf{V}^{view} , and $\mathbf{V}^{\text{light}}$ are calculated by SVD of the three Gram-like matrices:

$$\begin{aligned} \mathbf{G}^{\text{subj}} &= \mathbf{V}^{\text{subj}} \mathbf{S}^{\text{subj}^2} \mathbf{V}^{\text{subj}^T}, \\ \mathbf{G}^{\text{view}} &= \mathbf{V}^{\text{view}} \mathbf{S}^{\text{view}^2} \mathbf{V}^{\text{view}^T}, \\ \mathbf{G}^{\text{light}} &= \mathbf{V}^{\text{light}} \mathbf{S}^{\text{light}^2} \mathbf{V}^{\text{light}^T}. \end{aligned} \quad (9)$$

Then, $\mathbf{Z} \in \mathbb{R}^{m' \times n'_s n'_v n'_l}$ can be easily derived as

$$\mathbf{Z} = \mathbf{U}^T \mathbf{X} (\mathbf{V}^{\text{subj}} \otimes \mathbf{V}^{\text{view}} \otimes \mathbf{V}^{\text{light}}) \quad (10)$$

from (7). For a training image $\mathbf{x}_{s,v,l}$ assigned as one column of \mathbf{X} , the three factor parameters $\mathbf{v}_s^{\text{subj}}$, $\mathbf{v}_v^{\text{view}}$, and $\mathbf{v}_l^{\text{light}}$ are identical to the s th row of \mathbf{V}^{subj} , v th row of \mathbf{V}^{view} , and l th row of $\mathbf{V}^{\text{light}}$, respectively. In this paper, to solve for the three parameters of an arbitrary unlabeled image \mathbf{x} , one first calculates the Kronecker product of these parameters using (6):

$$\mathbf{v}^{\text{subj}} \otimes \mathbf{v}^{\text{view}} \otimes \mathbf{v}^{\text{light}} = \mathbf{Z}^+ \mathbf{U}^T \mathbf{x}, \quad (11)$$

where $^+$ denotes the Moore-Penrose pseudoinverse. To decompose the Kronecker product of multiple parameters into individual ones, two leading methods have been applied in [2] and [9]. The best rank-1 method [2] reshapes the vector $\mathbf{v}^{\text{subj}} \otimes \mathbf{v}^{\text{view}} \otimes \mathbf{v}^{\text{light}} \in \mathbb{R}^{n'_s n'_v n'_l}$ to the matrix $\mathbf{v}^{\text{subj}} (\mathbf{v}^{\text{view}} \otimes \mathbf{v}^{\text{light}})^T \in \mathbb{R}^{n'_s \times n'_v n'_l}$, and using SVD of this matrix, \mathbf{v}^{subj} is calculated as the left singular vector corresponding to the largest singular value. Another method is the rank- $(1, 1, \dots, 1)$ approximation using the alternating least squares method proposed in [9]. In this paper, we employed the decomposition method proposed in [2], which produced slightly better performances for face recognition than the method proposed in [9].

Based on the observation that the Gram-like matrices in (8) are formulated using the dot products, Multifactor Kernel PCA (MKPCA), a kernel-based extension of MPCA, was introduced [10]. If we define a kernel function k , the kernel versions of the Gram-like matrices in (8) can be directly calculated. Thus, for training images, \mathbf{V}^{subj} , \mathbf{V}^{view} , and $\mathbf{V}^{\text{light}}$ can be also calculated using eigen decomposition of these matrices. Equations (10) and (11) show that in order to obtain \mathbf{v}^{subj} , \mathbf{v}^{view} , and $\mathbf{v}^{\text{light}}$ for any test image, also called an out-of-sample image, \mathbf{x} , we must be able to calculate $\mathbf{U}^T \mathbf{X}$ and $\mathbf{U}^T \mathbf{x}$. Note that $\mathbf{U}^T \mathbf{X}$ and $\mathbf{U}^T \mathbf{x}$ are projections of training samples and a test sample onto nonlinear subspace, respectively, and these can be calculated by KPCA as shown in [11].

2.2. Linear Discriminant Analysis. Since Linear Discriminant Analysis (LDA) [3, 4] is a supervised learning algorithm, class labels of all samples are provided to the traditional LDA approach. Let $l_i \in 1, 2, \dots, c$ be the class label corresponding to \mathbf{x}_i , where $i = 1, 2, \dots, n$ and c is the number of classes. Let n_i be the number of samples in the class i such that $\sum_{i=1}^c n_i = n$. LDA calculates the optimal projection direction \mathbf{w} maximizing Fisher's criterion

$$J(\mathbf{w}) = \frac{\mathbf{w}^T \mathbf{S}_b \mathbf{w}}{\mathbf{w}^T \mathbf{S}_w \mathbf{w}}, \quad (12)$$

where \mathbf{S}_b and \mathbf{S}_w are the between-class and within-class scatter matrices:

$$\begin{aligned} \mathbf{S}_b &= \sum_{i=1}^c n_i (\mathbf{m}_i - \mathbf{m})(\mathbf{m}_i - \mathbf{m})^T, \\ \mathbf{S}_w &= \sum_{i=1}^n (\mathbf{x}_i - \mathbf{m}_i)(\mathbf{x}_i - \mathbf{m}_i)^T, \end{aligned} \quad (13)$$

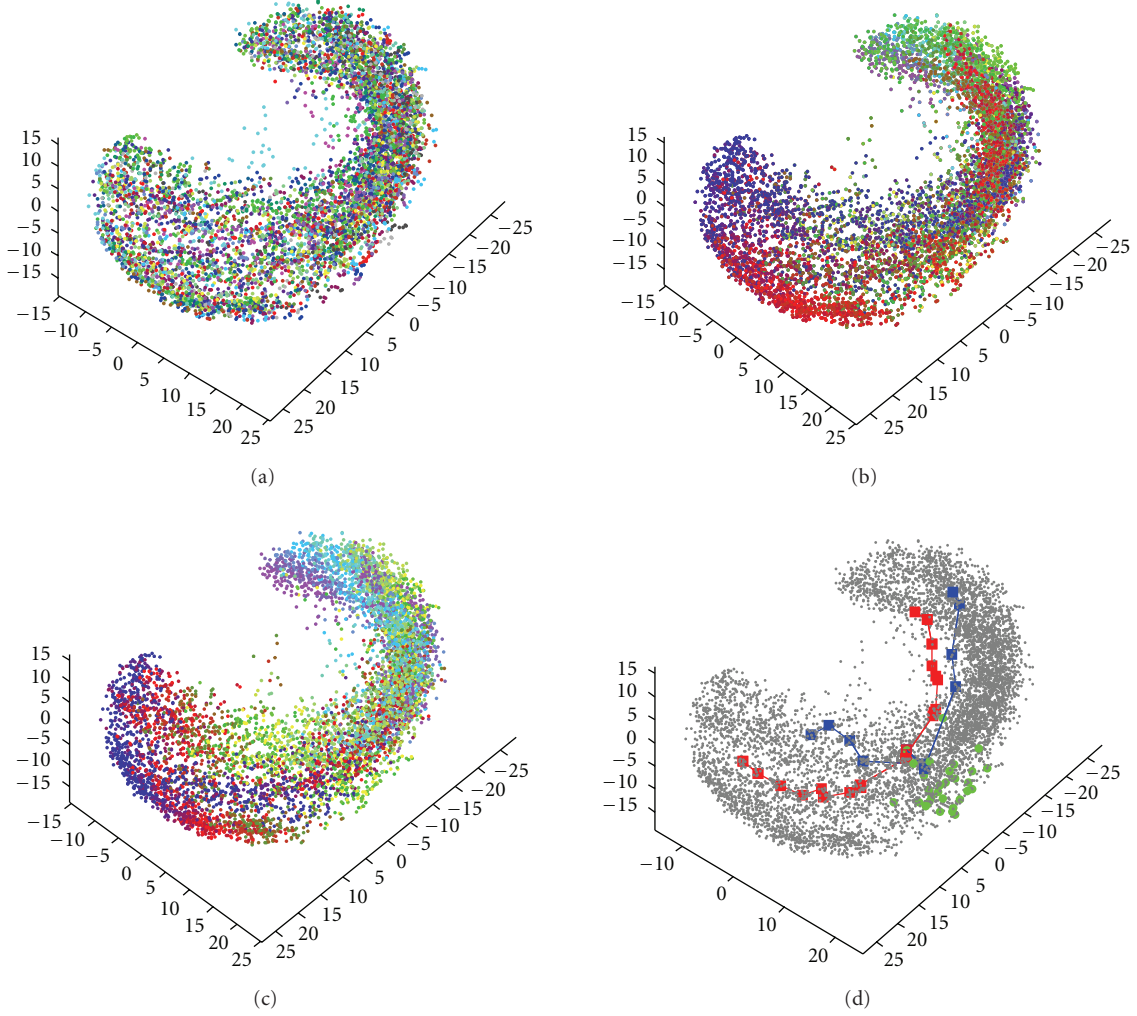


FIGURE 1: Low-dimensional representations of training images obtained by PCA using the CMU PIE database. (a) Each set of samples with the same color represents each subject's face images. (b) Each set of samples with the same color represents face images under each viewpoint. (c) Each set of samples with the same color represents face images under each lighting condition. (d) The red C-shape curve connects face images under various lighting conditions for one person and one viewpoint. The blue V-shape curve connects face images under various viewpoints for one person and one lighting condition. Green dots represent 30 subjects' face images under one viewpoint and one lighting condition. We can see that varying viewpoints and lighting conditions create clusters, rather than varying subjects.

where \mathbf{m}_i denotes the sample mean for the class i . The solution of (12) is calculated as the eigenvectors corresponding to the largest eigenvalues of the following generalized eigenvector problem:

$$\mathbf{S}_b \mathbf{w} = \lambda \mathbf{S}_w \mathbf{w}. \quad (14)$$

Since \mathbf{S}_w does not have full column rank and thus is not invertible, (14) can be solved not by eigen decomposition but instead by a generalized eigenvector problem. LDA obtains a low-dimensional representation \mathbf{y}_{LDA} for an arbitrary sample \mathbf{x} :

$$\mathbf{y}_{\text{LDA}} = \mathbf{W}^T \mathbf{x}, \quad (15)$$

where the columns of the matrix $\mathbf{W} \in \mathbb{R}^{n_p \times n'_p}$ consist of $\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_{p'}$. In other words, \mathbf{y}_{LDA} is the projection of \mathbf{x} onto the linear subspace spanned by $\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_{p'}$. Note

that $p' < c$. Despite the success of the LDA algorithm in many applications, the dimension of $\mathbf{y}_{\text{LDA}} \in \mathbb{R}^{n'_p}$ is often insufficient for representing each sample. This is caused by the fact that the number of available projection directions is lower than the class number c . To improve this limitation of LDA, variants of LDA, such as the null subspace algorithm [12] and a direct LDA algorithm [13], were proposed.

3. Limitations of Multifactor Analysis and Discriminant Analysis

LDA and MPCA have different advantages and disadvantages, which result from the fact that each method assumes different characteristics for data distributions. MPCA's subject parameters represent the average positions of a group of subjects across varying viewpoints and lighting conditions.

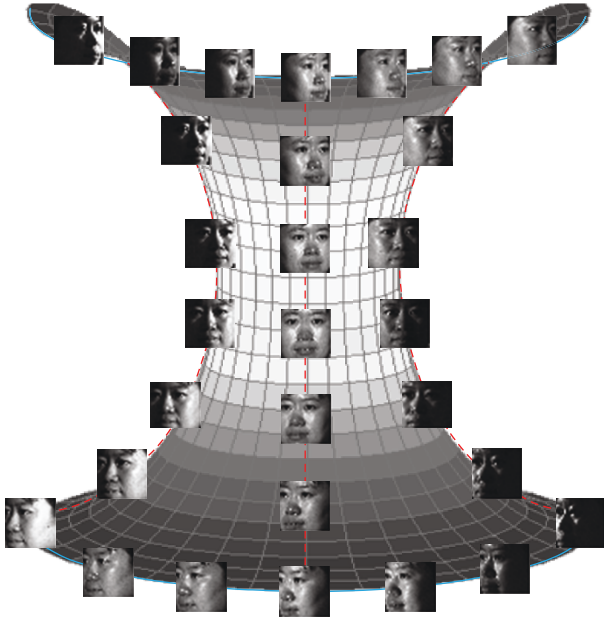


FIGURE 2: Ideal factor-specific submanifolds in an entire manifold on which face images lie. Each red curve connects face images only due to varying viewpoint while each blue curve connects face images only due to varying illumination.

MPCA's averaging is premised on the assumption that these subjects maintain similar relative positions in a data space under each viewpoint and lighting condition. On the other hand, LDA is based on the assumption that the samples of each class approximately create a Gaussian distribution. Thus, we can expect that the comparative performances of MPCA and LDA vary with the characteristics of a data set. For classification tasks, LDA sometimes outperforms MPCA; at other times MPCA outperforms LDA. In this section, we demonstrate the assumptions on which each method is based and the conditions where one can outperform the other.

3.1. The Assumption of LDA: Clusters Caused by Different Classes. Face recognition is a task to classify face images with respect to different subjects. LDA assumes that each class, that is, each subject, approximately causes a Gaussian distribution in a data set. Based on this assumption, LDA calculates a global linear subspace which is applied to the entire data set. However, a real-world face image set often includes other factors, such as viewpoints or lighting conditions in addition to differences between subjects. Unfortunately, the variation of viewpoints or lighting conditions often constructs global clusters across the entire data set while the variation of subjects creates only local distribution as shown in Figure 1. In the CMU PIE database, both viewpoints and lighting conditions create global clusters, as shown in Figures 1(b) and Figure 1(c), while a group of subjects creates a local distribution, as shown in Figure 1(a). Therefore, low-dimensional projections obtained by LDA are not appropriate for face recognition in these samples, which are not globally separable.

LDA inspires multiple advanced variants such as Kernel Discriminant Analysis (KDA) [14, 15], which can obtain nonlinear subspaces. However, these subspaces are still based on the analysis of the clusters distributed in a global data space. Thus, there is no guarantee that KDA can be successful if face images which belong to the same subject are scattered rather than distributed as clusters. In sum, LDA cannot be successfully applied unless, in a given data set, data samples are distributed as clusters due to different classes.

3.2. The Assumption of MPCA: Repeated Distributions Caused by Varying One Factor. MPCA is based on the assumption that the variation of one factor repeats similar shapes of distributions, and these common shapes rarely depend on the variation of other factors. For example, the subject parameters represent the averages of the relative positions of subjects in the data space across varying viewpoints and lighting conditions. To illustrate this, we consider viewpoint- and lighting-invariant subsets of a given face image set; each subset consists of the face images of n_s subjects captured under fixed viewpoint and lighting:

$$\mathbf{X}_{:,v,l} = [\mathbf{x}_{1,v,l} \ \mathbf{x}_{2,v,l} \ \cdots \ \mathbf{x}_{n_s,v,l}] \in \mathbb{R}^{n_p \times n_s} \quad (16)$$

That is, each column of $\mathbf{X}_{:,v,l}$ represents each image in this subset. As shown in Figure 4(a), there are $n_v n_l$ viewpoint- and lighting-invariant subsets, and \mathbf{G}^{subj} in (8) can be rewritten as the average of the Gram matrices calculated in these subsets:

$$\mathbf{G}^{\text{subj}} = \frac{1}{n_v n_l} \sum_{v=1}^{n_v} \sum_{l=1}^{n_l} \mathbf{X}_{:,v,l}^T \mathbf{X}_{:,v,l}. \quad (17)$$

In Euclidean geometry, the dot product between two vectors formulates the distance and linear similarity between them. Equation (9) shows that \mathbf{G}^{subj} is also the Gram matrix of a set of the column vectors of the matrix $\mathbf{S}^{\text{subj}} \mathbf{V}^{\text{subj}^T} \in \mathbb{R}^{n'_s \times n_s}$. Thus, these n_s column vectors represent the average distances between pairs of n_s subjects. Therefore, the row vectors of \mathbf{V}^{subj} , that is, the subject parameters, depend on these average distances between n_s subject across varying viewpoints and lighting conditions. Similarly, the viewpoint parameters and the lighting parameters depend on the average distances between n_v viewpoints and n_l lighting conditions, respectively, in a data space.

Figure 2 illustrates an ideal case to which MPCA can be successfully applied. Face images lie on a manifold, and viewpoint- and lighting-invariant subsets construct red and blue curves, respectively. Each red curve connects face images only due to varying illumination while each blue curve connects face images only due to varying viewpoints. Since all of the red curves have identical shapes, n_l different lighting conditions can be perfectly represented by n_l row vectors of $\mathbf{V}^{\text{light}} \in \mathbb{R}^{n_l \times n'_l}$. Also, since all of the blue curves have identical shapes, n_v different viewpoints can be perfectly represented by n_v row vectors of $\mathbf{V}^{\text{view}} \in \mathbb{R}^{n_v \times n'_v}$. For each factor, when these subsets construct similar structures with small variations, the average of these structures can successfully cover each sample.

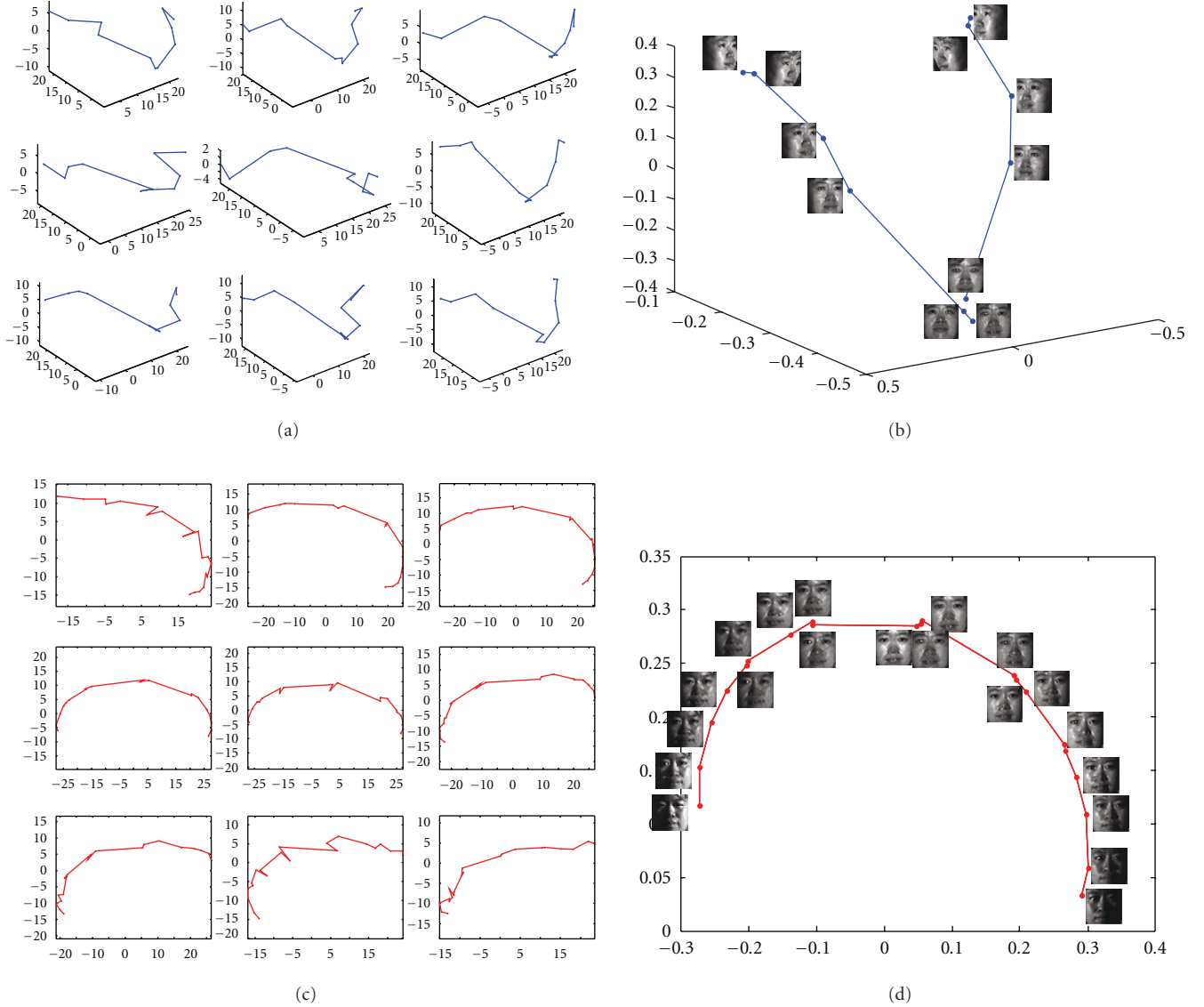


FIGURE 3: Low-dimensional representations of training images obtained by PCA and MPCA. (a) the PCA projections of 9 subjects' face images generated by varying viewpoints under one lighting condition. (b) the viewpoint parameters obtained by MPCA. (c) the PCA projections of 9 subjects' face images generated by varying lighting conditions under one viewpoint. (d) the lighting parameters obtained by MPCA.

We observe that each blue curve in Figure 3(a) that represents viewpoint variation seems to repeat a similar V-shape for each person and each lighting condition. Also, Figure 3(b) visualizes the viewpoint parameters \mathbf{y}_v , learned by MPCA; the curve connecting the viewpoint parameters roughly fits the average shape of the blue curves. As a result, \mathbf{y}_v in Figure 3(b) also has a V-shape. Also, the 3D visualization of the lighting parameters in Figure 3(d) roughly averages the C-shapes of red curves shown in Figure 3(c), each connecting face images under various lighting conditions for one person and one viewpoint. Similar observations were illustrated in [9].

Based on the above expectations, if varying just one factor generates dissimilar shapes of distribution, multilinear subspaces based on these average shapes do not represent

a variety of data distributions. In Figure 3(a), some curves have W-shapes while most of the other curves have V-shapes. Thus, in this case, we cannot expect reliable performances from MPCA because the average shape obtained by MPCA for each factor insufficiently covers individual shapes of curves.

4. Multifactor Discriminant Analysis

As shown in Section 3.1, for face recognition, LDA is preferred if in a given data set, face images are distributed as clusters due to different subjects. Unlike LDA, as shown in Section 3.2, MPCA can be successfully applied to face recognition if various subjects' face images repeat similar shapes of distributions under each viewpoint and lighting,

even if these subjects do not seem to create these clusters. In this paper, we propose a novel method which can offer the advantages of both methods. Our proposed method is based on an extension of LDA to multiple factor frameworks. Thus, we can call our method Multifactor Discriminant Analysis (MDA). From \mathbf{y}_{LDA} , MDA aims to remove the remaining characteristics which are caused by other factors, such as viewpoints and lighting conditions.

We start with the observation that MPCA is based on the relationships between \mathbf{y}_{PCA} , low-dimensional representations obtained by PCA, and multiple factor-specific parameters. Combining (3) and (7), we can see that the matrix $\mathbf{Y}_{\text{PCA}} \in \mathbb{R}^{n_p \times n_s n_v n_l}$ is rewritten as

$$\mathbf{Y}_{\text{PCA}} = \mathbf{U}^T \mathbf{X} = \mathbf{Z} (\mathbf{V}^{\text{subj}} \otimes \mathbf{V}^{\text{view}} \otimes \mathbf{V}^{\text{light}})^T. \quad (18)$$

Similarly, combining (2) and (7), for an arbitrary image \mathbf{x} , \mathbf{y}_{PCA} can be decomposed into three vectors by MPCA:

$$\mathbf{y}_{\text{PCA}} = \mathbf{U}^T \mathbf{x} = \mathbf{Z} (\mathbf{v}^{\text{subj}} \otimes \mathbf{v}^{\text{view}} \otimes \mathbf{v}^{\text{light}})^T \quad (19)$$

where \mathbf{y}_{PCA} is the low-dimensional representation of \mathbf{x} obtained by PCA. Thus, we can think that \mathbf{Z} performs a linear transformation which maps the Kronecker product of multiple factor-specific parameters to the low-dimensional representations provided by PCA. In other words, \mathbf{y}_{PCA} is decomposed into \mathbf{v}^{subj} , \mathbf{v}^{view} , and $\mathbf{v}^{\text{light}}$ by using the transformation matrix \mathbf{Z} .

In this paper, instead of decomposing \mathbf{y}_{PCA} , decomposing \mathbf{y}_{LDA} is proposed, where \mathbf{y}_{LDA} is the low-dimensional representation of \mathbf{x} provided by LDA, as defined in (15). \mathbf{y}_{LDA} often has more discriminant power than \mathbf{y}_{PCA} , but it still has the combined characteristics caused by multiple factors. Thus, we first formulate \mathbf{y}_{LDA} into the Kronecker product of the subject, viewpoint, and lighting parameters:

$$\mathbf{y}_{\text{LDA}} = \mathbf{W}^T \mathbf{x} = \mathbf{Z}' (\mathbf{v}^{\text{subj}'} \otimes \mathbf{v}^{\text{view}'} \otimes \mathbf{v}^{\text{light}'})^T, \quad (20)$$

where $\mathbf{W} \in \mathbb{R}^{n_p \times n'_p}$ is the LDA transformation matrix defined in (14) and (15). As reviewed in Section 2.2, n'_p , the number of available projection directions, is lower than the class number n_s ; $n'_p < n_s$. Note that \mathbf{y}_{LDA} in (20) is formulated in a similar way to \mathbf{y}_{PCA} in (19) using different factor-specific parameters and \mathbf{Z} . We expect $\mathbf{v}^{\text{subj}'}$ in (20), the subject parameter obtained by MDA, to be more reliable than both \mathbf{y}_{LDA} and \mathbf{v}^{subj} since $\mathbf{v}^{\text{subj}'}$ provides the advantages of the virtues of both LDA and MPCA. Using (15), we also calculate the matrix $\mathbf{Y}_{\text{LDA}} \in \mathbb{R}^{n'_p \times n_s n_v n_l}$ whose columns are the LDA projections of training samples.

While MPCA decomposes the data matrix $\mathbf{X} \in \mathbb{R}^{n_p \times n_s n_v n_l}$ consisting of training samples, our proposed MDA aims to decompose the LDA projection matrix \mathbf{Y}_{LDA} :

$$\mathbf{Y}_{\text{LDA}} = \mathbf{W}^T \mathbf{X} = \mathbf{Z}' (\mathbf{V}^{\text{subj}'} \otimes \mathbf{V}^{\text{view}'} \otimes \mathbf{V}^{\text{light}'})^T. \quad (21)$$

To obtain the factor-specific parameters of an arbitrary test image \mathbf{x} , we perform the following steps. During training, we first calculate the three orthogonal matrices, $\mathbf{V}^{\text{subj}'}$, $\mathbf{V}^{\text{view}'}$,

and $\mathbf{V}^{\text{light}'}$, and subsequently \mathbf{Z}' . Then, during testing, for the LDA projection \mathbf{y}_{LDA} of an arbitrary test image, we calculate the factor-specific parameters by decomposing $\mathbf{Z}'^+ \mathbf{y}_{\text{LDA}}$.

In Section 3.2, factor-specific parameters obtained by MPCA preserve the three Gram-like matrices \mathbf{G}^{subj} , \mathbf{G}^{view} , and $\mathbf{G}^{\text{light}}$ defined in (8). Figure 4 demonstrates that MPCA calculates subject, viewpoint, and lighting parameters using only the colored parts in the Gram matrix. These colored parts represent the dot products between pairs of samples that have only one varying factor. For example, the colored parts in Figure 4(a) represent the dot products of different subjects' face images under fixed viewpoint and lighting condition. Based on these observations, among the dot products of pairs of LDA projections, we only use the dot products which correspond to the colored parts of \mathbf{G} in Figure 4. Replacing \mathbf{x} with \mathbf{y}_{LDA} , we define three new Gram-like matrices, $\mathbf{G}^{\text{subj}'}$, $\mathbf{G}^{\text{view}'}$, and $\mathbf{G}^{\text{light}'}$:

$$\begin{aligned} [\mathbf{G}^{\text{subj}'}]_{m,n} &= \sum_{v=1}^{n_v} \sum_{l=1}^{n_l} \mathbf{y}_{\text{LDA},m,v,l}^T \mathbf{y}_{\text{LDA},n,v,l} \\ &= \sum_{v=1}^{n_v} \sum_{l=1}^{n_l} \mathbf{x}_{m,v,l}^T \mathbf{W} \mathbf{W}^T \mathbf{x}_{n,v,l} \end{aligned} \quad (22)$$

$$[\mathbf{G}^{\text{view}'}]_{m,n} = \sum_{s=1}^{n_s} \sum_{l=1}^{n_l} \mathbf{y}_{\text{LDA},s,m,l}^T \mathbf{y}_{\text{LDA},s,n,l}$$

$$[\mathbf{G}^{\text{light}'}]_{m,n} = \sum_{s=1}^{n_s} \sum_{v=1}^{n_v} \mathbf{y}_{\text{LDA},s,v,m}^T \mathbf{y}_{\text{LDA},s,v,n}$$

where $\mathbf{y}_{\text{LDA},s,v,l}$ denotes the LDA projection of a training image $\mathbf{x}_{s,v,l}$ of the s th subject under the v th viewpoint and the l th lighting condition. In (9), for MPCA, \mathbf{V}^{subj} , \mathbf{V}^{view} , and $\mathbf{V}^{\text{light}}$ are calculated as the eigenvector matrices of \mathbf{G}^{subj} , \mathbf{G}^{view} , and $\mathbf{G}^{\text{light}}$, respectively. In similar ways, for MDA, $\mathbf{V}^{\text{subj}'}$ $\in \mathbb{R}^{n_s \times n'_s}$, $\mathbf{V}^{\text{view}'}$ $\in \mathbb{R}^{n_v \times n'_v}$, and $\mathbf{V}^{\text{light}'}$ $\in \mathbb{R}^{n_l \times n'_l}$ can be calculated as the eigenvector matrices of $\mathbf{G}^{\text{subj}'}$, $\mathbf{G}^{\text{view}'}$, and $\mathbf{G}^{\text{light}'}$, respectively. Again, each row vector of $\mathbf{V}^{\text{subj}'}$ represents the subject parameter of each subject in a training set.

We remember that $\mathbf{Y}_{\text{LDA}} \in \mathbb{R}^{n'_p \times n_s n_v n_l}$ and $n'_p < n_s$. Thus, if we define the Gram matrix \mathbf{G}' as

$$\mathbf{G}' = \mathbf{Y}_{\text{LDA}}^T \mathbf{Y}_{\text{LDA}} = \mathbf{X}^T \mathbf{W} \mathbf{W}^T \mathbf{X}, \quad (23)$$

this matrix $\mathbf{G}' \in \mathbb{R}^{n_s n_v n_l \times n_s n_v n_l}$ does not have full column rank. If \mathbf{G}' is decomposed by SVD, \mathbf{G}' has $n_s - 1$ nonzero singular values at most. However, each of the matrices $\mathbf{G}^{\text{subj}'}$, $\mathbf{G}^{\text{view}'}$, and $\mathbf{G}^{\text{light}'}$ has full column rank since these matrices are defined in terms of the averages of different parts of \mathbf{G}' as shown in Figure 4. Thus, even if $n'_p < n_v$ or $n'_p < n_l$, one can calculate valid n_s , n_v , and n_l eigenvectors from $\mathbf{G}^{\text{subj}'}$, $\mathbf{G}^{\text{view}'}$, and $\mathbf{G}^{\text{light}'}$, respectively.

After calculating these three eigenvector matrices, $\mathbf{Z}' \in \mathbb{R}^{n'_p \times n_s n_v n_l}$ can be easily calculated as

$$\mathbf{Z}' = \mathbf{Y}_{\text{LDA}} (\mathbf{V}^{\text{subj}'} \otimes \mathbf{V}^{\text{view}'} \otimes \mathbf{V}^{\text{light}'}). \quad (24)$$

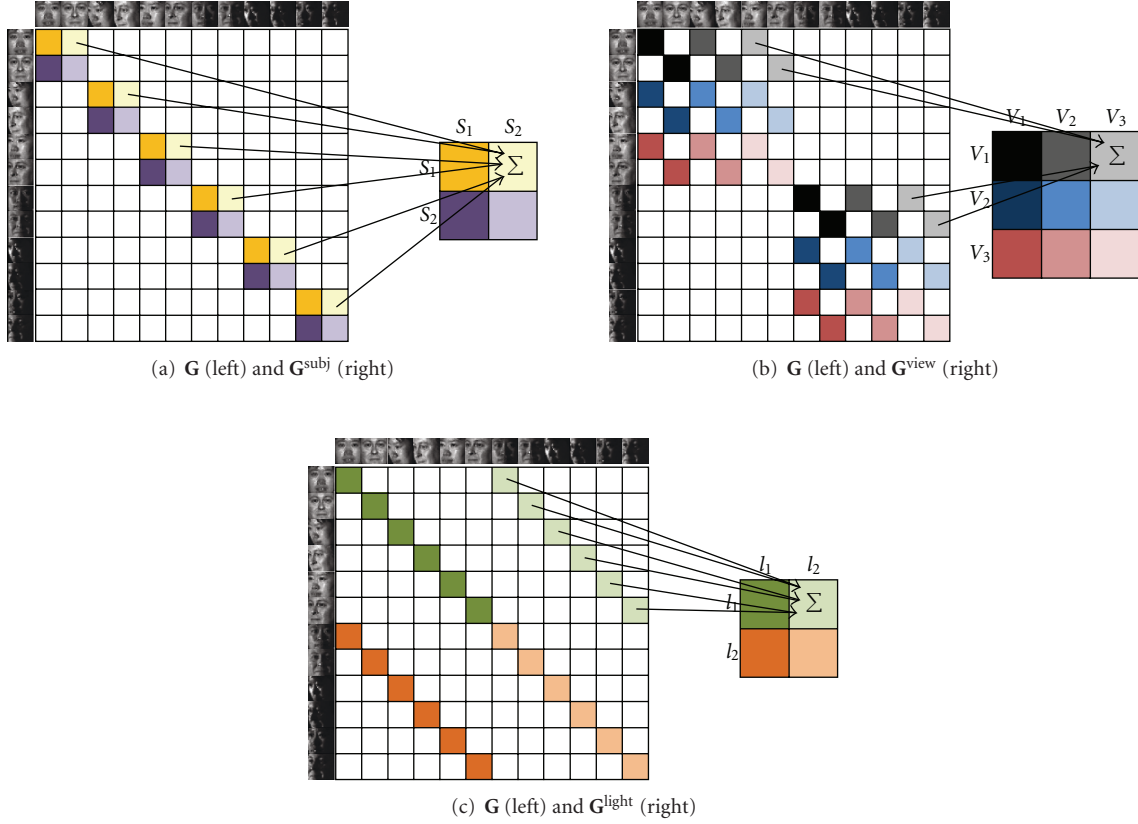


FIGURE 4: The relationships between the Gram matrix \mathbf{G} defined in (4) and each of the Gram-like matrices \mathbf{G}^{subj} , \mathbf{G}^{view} , and $\mathbf{G}^{\text{light}}$ defined in (8), where a training set has two subjects, three viewpoints, and two lighting conditions. Each of \mathbf{G}^{subj} , \mathbf{G}^{view} , and $\mathbf{G}^{\text{light}}$ is calculated as the average of parts of the Gram matrix \mathbf{G} . Each entry of these three Gram-like matrices is the average of same-color entries of \mathbf{G} . (a) \mathbf{G}^{subj} consists of averages of dot products which represent the averages of the pairwise relationships between a group of subjects. (b) \mathbf{G}^{view} consists of averages of dot products which represent the averages of the pairwise relationships between different viewpoints. (c) $\mathbf{G}^{\text{light}}$ consists of averages of dot products which represent the averages of the pairwise relationships between different lighting conditions.

Thus, using this transformation matrix \mathbf{Z}' , the Kronecker product of the three factor-specific parameters is calculated as

$$\mathbf{v}^{\text{subj}'} \otimes \mathbf{v}^{\text{view}'} \otimes \mathbf{v}^{\text{light}'} = \mathbf{Z}'^+ \mathbf{y}_{\text{LDA}}. \quad (25)$$

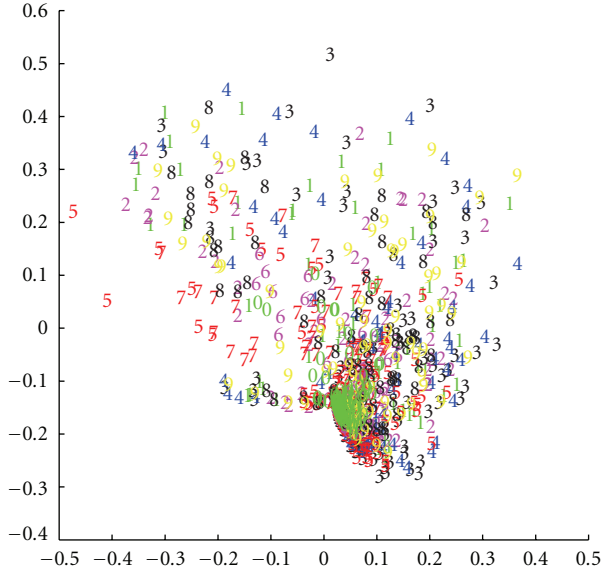
Again, as done in (11), by SVD of the matrix $\mathbf{v}^{\text{subj}'} (\mathbf{v}^{\text{view}'} \otimes \mathbf{v}^{\text{light}'})^T$, $\mathbf{v}^{\text{subj}'}$ is calculated as the left singular vector corresponding to the largest singular value. Consequently, we can obtain $\mathbf{v}^{\text{subj}'}$ of an arbitrary image test \mathbf{x} .

5. Experimental Results

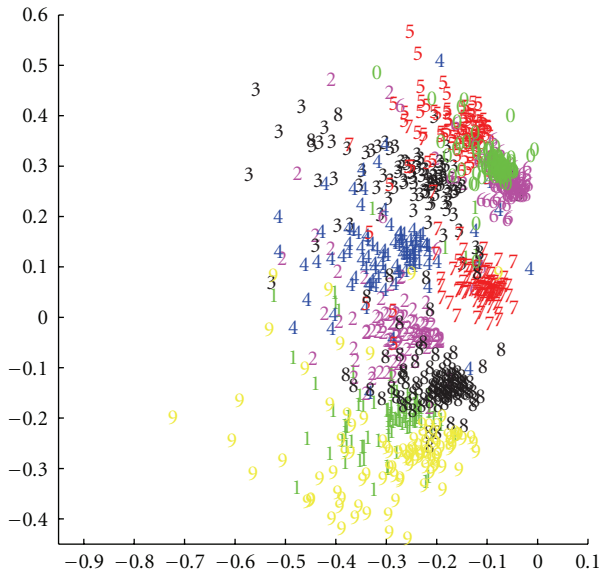
In this section, we demonstrate that Multifactor Discriminant Analysis is an appropriate method for dimension reduction of face images with varying factors. To test the quality of dimension reduction, we conducted face recognition tests. In all experiments, face images are aligned using eye coordinates and then cropped. Then, face images were resized to 32×32 gray-scale images, and each vectored image was normalized with unit norm and zero mean. After aligning and cropping, the left and right eyes are located at (9, 10) and (24, 10), respectively, in each 32×32 image.

For the face recognition experiments, we used two databases: the Extended YaleB database [16] and the CMU PIE database [17]. The Extended YaleB database contains 28 subjects captured under 64 different lighting conditions in 9 different viewpoints. For each of the subjects, we used all of the 9 viewpoints and the first 30 lighting conditions to reduce time for experiments. Among the face images, we used 10 lighting conditions in 5 viewpoints for each person for training and all of the remaining images for testing. Next, we used the CMU PIE database, which contains 68 individuals with 13 different viewpoints and 21 different lighting conditions. Again, to reduce time for experiments, we utilized 30 subjects. Also, we did not use two viewpoints: the leftmost profile and the rightmost profile. For each person, 5 lighting conditions in 5 viewpoints were used for training and all of the remaining images were used for testing. For each set of data, experiments were repeated 10 times using randomly selected lighting conditions and viewpoints. The averages of the results were reported in Tables 1 and 2.

We compare the performance of our proposed method, Multifactor Discriminant Analysis, and other traditional subspace projection methods with respect to dimension



(a)



(b)

FIGURE 5: Two dimensional projections of 10 classes in the Extended Yale B database. (a) features calculated by LDA, (b) subject parameters calculated by MDA.

reduction: PCA, MPCA, KPCA, and LDA. For PCA and KPCA, we used the subspaces consisting of the minimum numbers of eigenvectors whose cumulative energy is above 0.95. For MPCA, we set the threshold in pixel mode to 0.95 and the threshold in other modes to 1.0. KPCA used RBF kernels with σ set to 100. We compared the rank-1 recognition rates of all of the methods using the simple cosine distance.

As shown in Tables 1 and 2, our proposed method, Multifactor Discriminant Analysis, outperforms the other

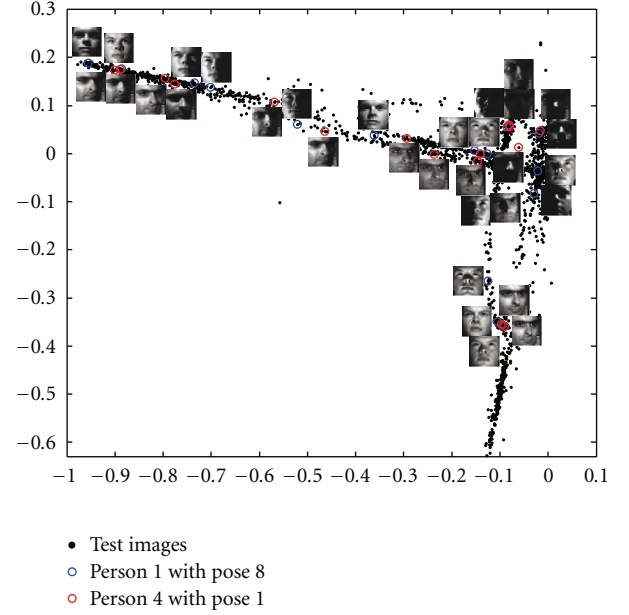


FIGURE 6: The first two coordinates of lighting feature vectors computed by Multifactor Discriminant Analysis using the Extended Yale database.

TABLE 1: Rank-1 recognition rate on the Extended YaleB database.

	untrained lighting	untrained viewpoints	Untrained viewpoints & lighting
PCA	0.87 ± 0.03	0.64 ± 0.03	0.59 ± 0.05
MPCA	0.90 ± 0.01	0.70 ± 0.05	0.65 ± 0.06
KPCA	0.88 ± 0.03	0.67 ± 0.04	0.64 ± 0.06
LDA	0.89 ± 0.03	0.65 ± 0.03	0.62 ± 0.05
MDA	0.94 ± 0.03	0.77 ± 0.04	0.70 ± 0.05

methods for face recognition. This seems to be because Multifactor Discriminant Analysis offers the combined virtues of both multifactor analysis methods and discriminant analysis methods. Like multilinear subspace methods, Multifactor Discriminant Analysis can analyze one sample in a multiple factor framework, which improves face recognition performance.

Figure 5 shows two dimensional projections of 10 subjects under varying viewpoints and lighting conditions calculated by LDA and Multifactor Discriminant Analysis. For each image, while LDA calculated one kind of projection vector as shown in Figure 5(a), Multifactor Discriminant Analysis obtained individual projection vectors for subjects, viewpoint and lighting. Among the factor parameters, Figure 5(b) shows subject parameters obtained by MDA. Since these parameters are independent from varying viewpoints and lighting conditions, the subject parameters of face images are distributed as clusters created by varying subjects rather than the scattered results in Figure 5(a). For the same reason, Tables 1 and 2 show that MPCA and Multifactor Discriminant Analysis outperformed PCA and LDA respectively.

TABLE 2: Rank 1 recognition rate on the CMU PIE database.

	untrained lighting	untrained viewpoints	untrained viewpoints & lighting
PCA	0.89 ± 0.06	0.70 ± 0.05	0.22 ± 0.05
MPCA	0.91 ± 0.04	0.74 ± 0.05	0.24 ± 0.06
KPCA	0.91 ± 0.04	0.73 ± 0.05	0.23 ± 0.06
LDA	0.90 ± 0.06	0.72 ± 0.05	0.23 ± 0.05
MDA	0.96 ± 0.04	0.79 ± 0.04	0.27 ± 0.06

Also, Figure 6 shows the first two coordinates of the lighting features calculated by Multifactor Discriminant Analysis for the face images of two different subjects in different viewpoints. These two-dimensional mappings are continuously distributed with steadily varying lighting while differences in subjects or viewpoint appear to be relatively insignificant. For example, for both Person 1 in Viewpoint 8 and Person 4 in Viewpoint 1, the mappings for face images that were lit from the subjects' right side appear on the top left-hand corner, while dark images appear on the top-right corner; images captured under neutral lighting conditions lie on the bottom right. On the other hand, any two images captured under similar lighting conditions tend to be located close to each other even if they are of different subjects in different viewpoints. Therefore, we can conclude that the lighting features calculated by our proposed MDA preserve neighbors for lighting, which are captured under similar lighting conditions.

6. Conclusion

In this paper, we propose a novel dimension reduction method for face recognition: Multifactor Discriminant Analysis. Multifactor Discriminant Analysis can be thought of as an extension of LDA to multiple factor frameworks providing both multifactor analysis and discriminant analysis. Moreover, we have shown through experiments that MDA extracts more reliable subject parameters compared to the low-dimensional projections obtained by LDA and MPCA. These subject parameters obtained by MDA represent locally repeated shapes of distributions due to differences in subjects for each combination of other factors. Consequently, MDA can offer more discriminant power, making full use of both global distribution of the entire data set and local factor-specific distribution. Reference [6] introduced another method which is theoretically based on both MPCA and LDA: Multilinear Discriminant Analysis. However, Multilinear Discriminant Analysis cannot analyze multiple factor frameworks, while our proposed Multifactor Discriminant Analysis can. Relevant examples are shown in Figure 5 where our proposed approach has been able to yield a discriminative two dimensional subspace that can cluster multiple subjects in the Yale-B database. On the other hand, LDA completely spreads the data samples into one global undiscriminative distribution of data samples. These results show the dimension reduction power of our approach in the presence of nuisance factors such as viewpoints and

lighting conditions. This improved dimension reduction power will allow us to have reduced size feature sets (optimal for template storage) and increased matching speed due to these smaller dimensional features. Our approach is thus attractive for robust face recognition for real-world defense and security applications. Future work will include evaluating this approach on larger data sets such as the CMU Multi-PIE database and NIST's FRGC and MBGC databases.

References

- [1] M. A. O. Vasilescu and D. Terzopoulos, "Multilinear image analysis for facial recognition," in *Proceedings of the International Conference on Pattern Recognition*, vol. 1, no. 2, pp. 511–514, 2002.
- [2] M. A. O. Vasilescu and D. Terzopoulos, "Multilinear independent components analysis," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, vol. 1, pp. 547–553, San Diego, Calif, USA, 2005.
- [3] K. Fukunaga, *Introduction to Statistical Pattern Recognition*, Academic Press, San Diego, Calif, USA, 2nd edition, 1999.
- [4] A. M. Martinez and A. C. Kak, "PCA versus LDA," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 23, no. 2, pp. 228–233, 2001.
- [5] M. Turk and A. Pentland, "Eigenfaces for recognition," *Journal of Cognitive Neuroscience*, vol. 3, no. 1, pp. 71–86, 1991.
- [6] S. Yan, D. Xu, Q. Yang, L. Zhang, X. Tang, and H.-J. Zhang, "Multilinear discriminant analysis for face recognition," *IEEE Transactions on Image Processing*, vol. 16, no. 1, pp. 212–220, 2007.
- [7] G. H. Golub and C. F. V. Loan, *Matrix Computations*, The Johns Hopkins University Press, London, UK, 1996.
- [8] L. De Lathauwer, B. De Moor, and J. Vandewalle, "A multilinear singular value decomposition," *SIAM Journal on Matrix Analysis and Applications*, vol. 21, no. 4, pp. 1253–1278, 2000.
- [9] M. A. O. Vasilescu and D. Terzopoulos, "Multilinear projection for appearance-based recognition in the tensor framework," in *Proceedings of the IEEE International Conference on Computer Vision (ICCV '07)*, pp. 1–8, 2007.
- [10] Y. Li, Y. Du, and X. Lin, "Kernel-based multifactor analysis for image synthesis and recognition," in *Proceedings of the IEEE International Conference on Computer Vision*, vol. 1, pp. 114–119, 2005.
- [11] B. Scholkopf, A. Smola, and K.-R. Muller, "Nonlinear component analysis as a kernel eigenvalue problem," in *Neural Computation*, pp. 1299–1319, 1996.
- [12] X. Wang and X. Tang, "Dual-space linear discriminant analysis for face recognition," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 564–569, 2004.
- [13] H. Yu and J. Yang, "A direct LDA algorithm for high dimensional data-with application to face recognition," *Pattern Recognition*, pp. 2067–2070, 2001.
- [14] G. Baudat and F. Anouar, "Generalized discriminant analysis using a kernel approach," *Neural Computation*, vol. 12, no. 10, pp. 2385–2404, 2000.
- [15] S. Mika, G. Ratsch, J. Weston, B. Scholkopf, and K.-R. Muller, "Fisher discriminant analysis with kernels," in *Proceedings of the IEEE Workshop on Neural Networks for Signal Processing*, pp. 41–48, 1999.

- [16] K.-C. Lee, J. Ho, and D. J. Kriegman, "Acquiring linear subspaces for face recognition under variable lighting," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, no. 5, pp. 684–698, 2005.
- [17] T. Sim, S. Baker, and M. Bsat, "The CMU pose, illumination, and expression database," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 25, no. 12, pp. 1615–1618, 2003.